Volume 319, number 3, 207–211 FEBS 12209 March 1993

Genetic polymorphism of cytochrome P4502E1 in a Swedish population

Relationship to incidence of lung cancer

Irene Persson[a], Inger Johansson[a], Helene Bergling[a], Marja-Liisa Dahl[d], Janeric Seidegård[e], Ragnar Rylander[c], Agneta Rannug[b], Johan Högberg[b] and Magnus Ingelman Sundberg[a]

[a]*Department of Physiological Chemistry, Karolinska Institutet, Stockholm, Sweden,* [b]*National Institute of Occupational Health, Solna, Sweden,* [c]*Department of Environmental Medicine, University of Gothenburg, Gothenburg, Sweden,* [d]*Department of Clinical Pharmacology, Huddinge University Hospital, Karolinska Institutet, Huddinge, Sweden* and [e]*Molecular Ecogenetics, The Wallenberg Laboratory and Astra Draco AB, Lund, Sweden*

Received 29 December 1992; revised version received 23 January 1993

Genetic polymorphism of *CYP2E1* was investigated among 195 Swedish patients with lung cancer and 206 controls. Three different polymorphic sites were found, all in introns, using RFLP and the restriction enzymes *Dra*I, *Rsa*I and *Taq*I. The frequencies of the rare alleles were 0.08–0.18 and much lower than previously described among Japanese. No significant difference in distribution of the polymorphic alleles between controls and lung cancer patients was evident, in contrast to results of a previous Japanese study. However, examination of a polymorphic site in the 5′-flanking region, within a putative binding motif for the hepatic transcription factor HNF-1, revealed a significantly less frequent distribution of the mutated allele (c_2) among the lung cancer patients as compared to controls. It is concluded that major interethnic differences exist in the genetic polymorphism of *CYP2E1* and that people carrying the c_2 allele might be at lower risk for developing lung cancer.

Lung cancer; Ethanol; Polymorphism; Drug metabolism

1. INTRODUCTION

Ethanol-inducible cytochrome P450 2E1 (CYP2E1) metabolizes several suspect as well as established carcinogens [1,2]. More than 75 specific substrates for CYP2E1 have been identified [2], most of them small and hydrophobic in nature. Among the substrates are several potentially important carcinogens such as benzene, dimethylnitrosoamine, vinyl chloride and diethyldithiocarbamate. CYP2E1 is localised in several different tissues, among them the brain [3] and the lung [4], but the highest expression is in the liver. The hepatic distribution of CYP2E1 is almost entirely in the centrilobular region [5] and the amount of the isozyme can here reach approximately 0.1 mM after induction. The regional distribution of CYP2E1 correlates well to the areas where the hepatotoxic chemicals, known to be activated by CYP2E1, exert their toxic action [2].

The *CYP2E1* gene has been shown to be polymorphically distributed among Japanese, as revealed from RFLP analyses using the restriction enzymes *Rsa*I and *Dra*I [6], and among Caucasians, as revealed by *Taq*I RFLP [7]. In addition, a polymorphic site in the 5′-flanking region constituting a putative HNF-1 binding site, has been described among Japanese [8,9]. Because of the important role of CYP2E1 in the metabolic activation of carcinogens, it might be speculated that different allelic *CYP2E1* forms would cause formation of altered variants of CYP2E1, alternatively cause different levels of enzyme expression. This might be important for the rate of metabolic activation and, hence, for the individual susceptibility to carcinogen exposure. In a recent study, where *Dra*I RFLP of *CYP2E1* was analyzed among 47 lung cancer patients and 56 controls, no individuals homozygous for an allelic form termed C were found in the cancer group, whereas 6 such individuals were present in the control group [10]. This difference was considered to be significant. These results prompted us to evaluate the frequency of subjects polymorphic for the *CYP2E1* gene in a caucasian population and evaluate its possible relationship to incidence of lung cancer.

Correspondence address: M. Ingelman-Sundberg, Department of Physiological Chemistry, Karolinska Institutet, PO Box 60 400, S-104 01 Stockholm, Sweden.

2. MATERIALS AND METHODS

2.1. *Isolation of genomic DNA*

Blood samples were obtained from 151 Swedish healthy controls, 55 bronchitis patients and 195 Swedish lung cancer patients from six different hospitals in the areas of Stockholm, Gothenburg, Malmö and Lund. The material from Gothenburg (40 cases and 16 controls collected from two hospitals in the West of Sweden) was from a larger prospective study. The lung cancer patients all had a history of smoking and the majority of the histological diagnoses included small cell cancer, large cell cancer and squamous cell carcinoma. 27 specimens were obtained from subjects with adenocarcinoma. DNA samples

from 136 healthy unrelated controls were obtained from sources previously described [11]. Genomic DNA was isolated using a guanidinium-isothiocyanate method. Blood from the cancer patients and the remaining controls was obtained frozen in heparin or EDTA-containing tubes. After thawing, the samples were subjected to lysis and proteinase digestion before extraction with phenol and chloroform [12]. DNA was isolated after ethanol precipitation.

2.2. *RFLP analysis*

Seven μg of genomic DNA was digested overnight with 28 U of the restriction endonuclease (Boehringer-Mannheim) in question and then subjected to electrophoresis in 1% agarose gel at 30 V for 24 h. Blotting was performed with Gene Screen Plus filters (NEN, DuPont) and the filter hybridised with a nick translated [^{32}P]dCTP (Amersham) labelled CYP2E1 cDNA probe (kindly provided by Dr. F.J. Gonzalez, NCI, NIH, Bethesda).

2.3. *Polymorphic analysis of the 5′-flanking region*

One μg of DNA was subjected to amplification using two designed primers; 5′-CCCGTGAGCCAGTCGAGT-3′ (−1,378 to −1,361) and 5′-ATACAGACCCTCTTCCAC-3′ (−869 to −886), in a reaction creating a 510 bp fragment containing both the *Rsa*I and *Pst*I polymorphic sites in the 5′-flanking area [8]. These polymorphisms have been reported to be linked to each other so that each allele type only contains one of these restriction sites in the 5′-flanking region [8]. After initial denaturation for 1.5 min at 94°C, the PCR reaction was carried out for 35 cycles under the following conditions: denaturation at 94°C for 1 min; annealing temperature, 52°C for 1min; extention period at 72°C for 1 min. Five μl of the product was subjected to electrophoresis on a 1.8% agarose gel. The remainder was diluted 3-fold with water and buffer and subsequently digested with *Rsa*I for 8–12 h. The results were analyzed as complete, incomplete or no fragment cleavage. The most common allelic form, c_1, has got a *Rsa*I site. When none, or incomplete cleavage was obtained, the whole procedure was repeated and the fragment digested with *Pst*I as a control for c_2.

2.4. *Statistical analysis*

The χ^2-test was used and $P < 0.05$ considered significant.

3. RESULTS

3.1. *RFLP-analysis of the* CYP2E1 *gene among Swedish controls and lung cancer cases*

Southern blot analysis of genomic DNA from healthy Swedish subjects revealed RFLPs with the restriction enzymes *Dra*I, *Rsa*I and *Taq*I (Fig. 1). The polymorphic fragment lengths obtained with *Dra*I were 4.1 and 4.7 kb. This is different from the length of the *Dra*I polymorphic fragments 5.5 and 4.1 kb, as originally described by Uematsu et al. [6]. In order to exclude interethnic differences on this specific point, *Dra*I RFLP was performed on DNA obtained from 6 healthy Japanese living in Stockholm (cf. Fig. 1). However, under our conditions the fragmentation pattern of the Japanese DNA was identical with that seen using the Swedish samples. These results indicate that the *Dra*I polymorphic site is located in intron 6 (Fig. 1) and not in intron 3 as originally proposed [10]. Based on these results and because of the presence of two neighbouring *Dra*I cleavage sites in intron 3 [13], we propose that the polymorphic *Dra*I site was erroneously identified in the previous report [10] and therefore use the same nomenclature (C and D) for the polymorphic variant alleles.

The frequencies of the different *CYP2E1* alleles identified by RFLP were evaluated using DNA obtained from healthy controls, bronchitis patients and from patients diagnosed for lung cancer (Table I). As compared to the previous reports [6,10], the frequency of the *Dra*I C-allele was much lower (0.10) in the Swedish populations studied and in total, only two subjects homozygous for this allele were found. No significant difference in the distribution of the C-allele between controls and cases could be seen.

*Taq*I and *Rsa*I RFLP-analysis revealed the presence of two different allelic variants in both cases (Fig. 1, Tables I and II). The frequency of the rare allele, detected with *Rsa*I, was also much lower among the Swedes as compared to Japanese. No difference between controls and cases was seen for the *Taq*I polymorphism.

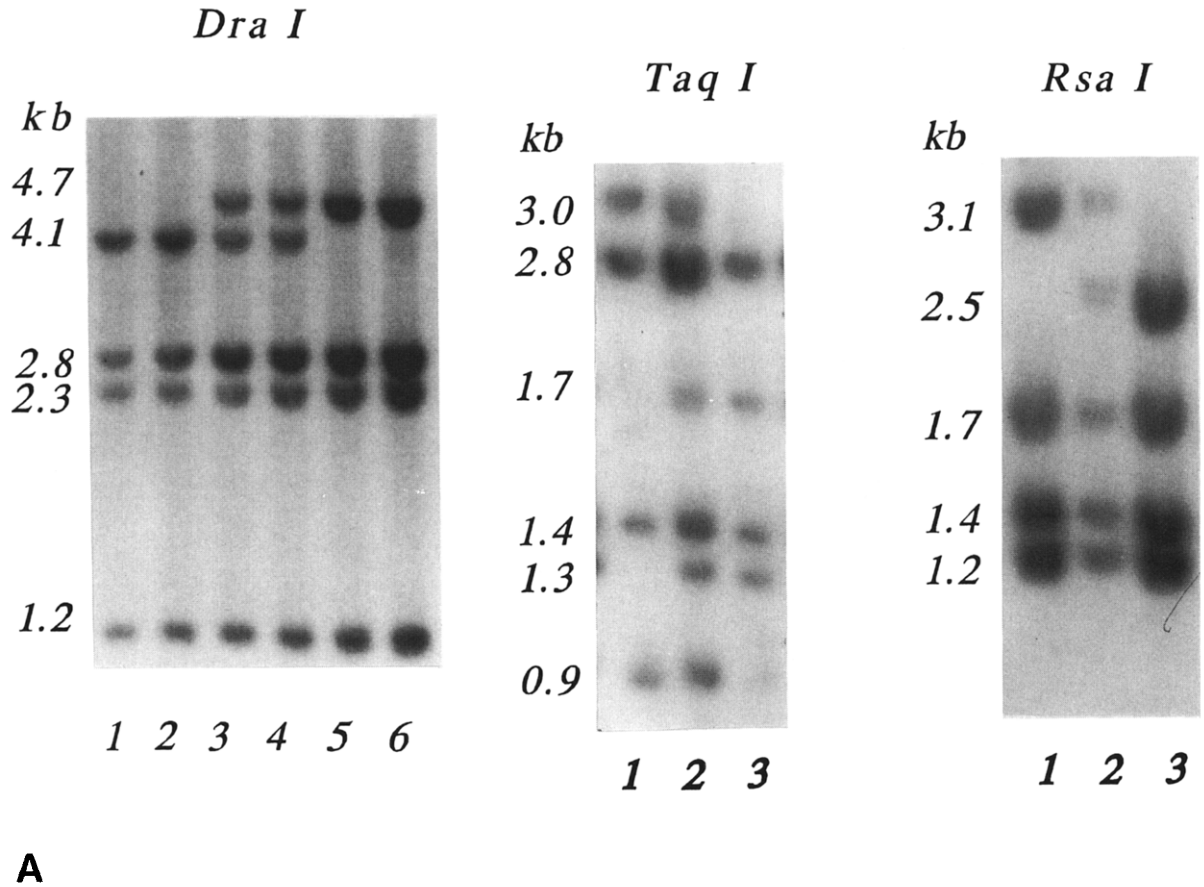
From Table II it is apparent that the *Taq*I polymorphism was unrelated to the polymorphisms detected with either *Dra*I or *Rsa*I, whereas the 92% of the *Dra*I D alleles were of the *Rsa*I H-type. Because of this we did not consider it of importance to investigate the distribution of the *Rsa*I polymorphism among lung cancer cases.

3.2. *Analysis of the 5′-flanking region*

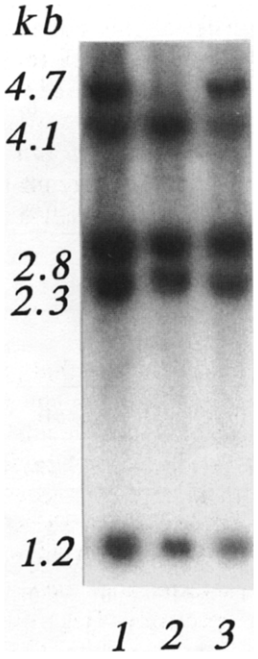
Previous data have revealed the existence of a polymorphism in the 5′-flanking region of *CYP2E1* at a location of −1,017 bp in a putative binding region for the transcription factor HNF-1 [8,9]. Preliminary data indicating the importance of this polymorphism for transcriptional efficiency have been given [9]. Analysis of our DNA samples for this polymorphism revealed that the frequency of the rare allele (c_2) was lower (0.05) among healthy Swedes as compared to Japanese. Interestingly, there was a significant difference ($P < 0.05$) between the two groups studied and among the lung cancer patients the corresponding allele frequency was only 0.02 (Table I). All haplotypes carrying the c_2 allele were of the *Dra*I C-type.

4. DISCUSSION

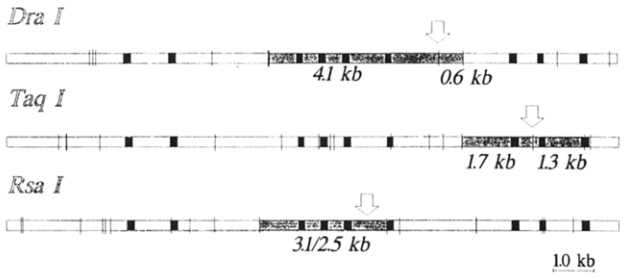
The data here presented show that pronounced interethnic differences are evident for the genetic polymorphism of *CYP2E1*. Thus, the rare alleles are much less prominent among Swedes, as compared to previous results obtained from Japanese subjects [6,9,10]. This is true for the polymorphic sites analyzed in the 5′-flanking region at −1,017 bp [13] and the polymorphic sites in introns 5 and 6, detected with the restriction enzymes *Rsa*I and *Dra*I, respectively. Regarding the polymorphic site in intron 7, as seen with *Taq*I, it is, however, evident that the frequency of the rare allele among Swedes is the same as previously seen among Americans [7]. These findings further emphasize the importance of taking interethnic differences into consideration when



A



B



C

Fig. 1. RFLP-analysis of the *CYP2E1* gene using the restriction enzymes *Dra*I, *Taq*I and *Rsa*I. (A) RFLPs obtained using Caucasian samples. (B) *Dra*I RFLP obtained using leucocytes from two Japanese subjects (lanes 1 and 2) as compared to a Swedish individual (lane 3). (C) Schematic drawing of the *CYP2E1* gene illustrating the restriction sites. The shadowed area indicates the polymorphic fragments obtained. A full-length cDNA probe was used in the Southern blot experiments.

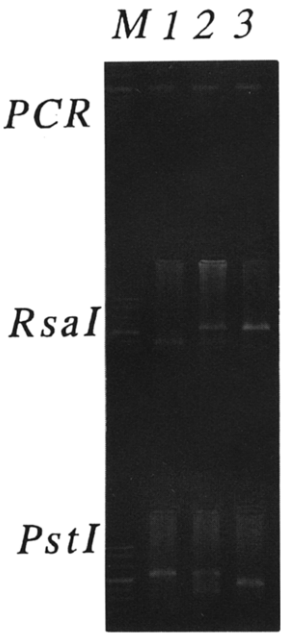


Fig. 2. PCR and RFLP analysis of polymorphism in the 5′-flanking region of the *CYP2E1* gene. A 510 bp fragment (−1,377 to −868) was amplified by PCR and subjected to cleavage using the restriction enzymes *Rsa*I or/and *Pst*I. The figures show the results obtained with DNA from subjects homozygous for the c_1 allele (lane 1), heterozygous (c_1/c_2, lane 2) and homozygous for c_2 (lane 3).

discussing specific allelic forms as genetic markers for cancer or disease. Previous results from our laboratory have revealed important interethnic differences in the structure of the *CYP2D* locus [14–16] with implications for the frequency of poor metabolizers of debrisoquine among Chinese as compared to Caucasians.

In contrast to previous results obtained by genetic analysis of Japanese subjects, we could not identify any significant difference in the distribution of the *Dra*I C-allele between controls and lung cancer patients (Table I). The same was true for polymorphic distribution of the allelic forms identified with RFLP and the restriction enzyme *Taq*I. In general, the frequencies of the rare *CYP2E1* alleles detected by these methods among the Swedes would be too low in order to allow to utilize them as any potentially valuable markers for lung cancer. Thus, there is a need for further research to find other mutations among the Caucasian *CYP2E1* gene that could fulfil such a requirement.

About 50% of the *Dra*I C-alleles also exhibited the −1,017 C → T mutation in the 5′-flanking region. Interestingly, there was no homogenous distribution of this mutation between the controls and the samples from lung cancer patients (Table I). Thus, the frequency of this mutation among lung cancers was only 40% compared to its presence in controls. This mutation has been assumed to have a functional role [9]. The allelic form c_2 does less effectively bind nuclear factors to the region of the mutation. A similar mutation has been observed in the *CYP2D6* gene among Chinese of the *Xba*I 44 kb haplotype and the distribution of this mutation very well correlates to a diminished capability for debrisoquine metabolism among the individuals (Johansson,

Table I

Genetic polymorphism of the *CYP2E1* gene in Swedish controls, healthy persons (A) and bronchitis patients (B), and lung cancer patients as revealed by RFLP

(A) *Dra*I RFLP

	Controls				Lung cancer patients	
	A (n=152)		B (n=54)		(n = 193)	
	n	(%)	n	(%)	n	(%)
Genotype						
DD	123	81	43	80	160	83
CD	28	18	10	18	33	17
CC	1	<1	1	2	0	0
Allele frequencies						
D		0.90		0.89		0.91
C		0.10		0.11		0.09

(B) *Taq*I RFLP

	Controls				Lung cancer patients	
	A (n=145)		B (n=52)		(n = 182)	
	n	(%)	n	(%)	n	(%)
Genotype						
A_2A_2	116	80	44	85	148	81
A_1A_2	28	19	8	15	34	19
A_1A_1	1	<1	0	0	0	0
Allele frequencies						
A_2		0.90		0.92		0.91
A_1		0.10		0.08		0.09

(C) 5′-flanking region

	Controls				Lung cancer patients	
	A (n=148)		B (n=54)		(n = 184)	
	n	(%)	n	(%)	n	(%)
Genotype						
c_1c_1	133	90	49	91	176	96
c_1c_2	14	9	5	9	8	4
c_2c_2	1	<1	0	0	0	0
Allele frequencies						
c_1		0.95		0.95		0.98
c_2		0.05		0.05		0.02*

*The frequency of the c_2 allele was significantly ($P < 0.05$) lower among the lung cancer cases as compared to the control groups.

Table II

Summary of the genetic polymorphisms of *CYP2E1* obtained using DNA from healthy Swedes

		*Dra*I		
		DD (n = 104)	CD (n = 25)	CC (n = 1)
*Rsa*I	HH	87	–	–
	HI	17	22	–
	II	–	3	1
*Taq*I	A_2A_2	79	20	1
	A_1A_2	24	5	–
	A_1A_1	1	–	–
5′-flanking	c_1c_1	104	13	–
	c_1c_2	–	12	–
	c_2c_2	–	–	1

Ying, Bertilsson and Ingelman-Sundberg, unpublished observation). It may be suggested that the *CYP2E1* gene in vivo is less efficiently expressed among people carrying the c_2 allele. The data here presented furthermore indicate these subjects may be at lower risk for development of smoke-induced lung cancer. The basis for the detailed functional relationship has, however, to await further research.

After completion of this study, Kato et al. [17] published that there is an interethnic difference in the distribution of the two mutations in the 5′-flanking region of the c_2 allele and that these two mutations are not in total linkage among Japanese and African–Americans. Furthermore, they found no significant difference in frequency of this allele between 67 lung cancer patients and 122 controls.

Acknowledgements: This work was supported by grants from the Swedish Environment Protection Fund and from the Swedish Cancer Society.

REFERENCES

[1] Guengerich, P.F., Kim, D.-H. and Iwasaki, M. (1991) Chem. Res. Toxicol. 4, 168–179.
[2] Terelius, Y., Lindros, K.O., Albano, E. and Ingelman-Sundberg, M. (1992) Frontiers of Biotransformation 8, 187–232.
[3] Hansson, T., Tindberg, N., Ingelman-Sundberg, M. and Köhler, C. (1990) Neuroscience 34, 451–463.
[4] Tindberg, N. and Ingelman-Sundberg, M. (1989) Biochemistry 28, 4499–4504.
[5] Ingelman-Sundberg, M., Johansson, I., Penttilä, K.E., Glaumann, H. and Lindros, K.O. (1988) Biochem. Biophys. Res. Commun. 157, 55–60.
[6] Uematsu, F., Kikuchi, H., Ohmachi, T., Sagami, I., Motomiya, M., Kamataki, T., Komori, M. and Watanabe, M. (1991) Nucleic Acids Res. 19, 2803.
[7] McBride, O.W., Umeno, W., Gelboin, H.V. and Gonzalez, F.J. (1987) Nucleic Acids Res. 15, 10071.
[8] Watanabe, J., Hayashi, S.-I., Nakachi, K., Imai, K., Suda, Y., Sekine, T. and Kawajiri, K. (1990) Nucleic Acids Res. 18, 7194.
[9] Hayashi, S.-I., Watanabe, J. and Kawajiri, K. (1991) J. Biochem. 110, 559–565.
[10] Uematsu, F., Kikuchi, H., Motomiya, M., Tatsuya, A., Sagami, I., Ohmachi, T., Wakui, A., Kanamaru, R. and Watanabe, M. (1991) Jpn. J. Cancer Res. 82, 254–256.
[11] Dahl, M.-L., Johansson, I., Porsmyr-Palmertz, M., Ingelman-Sundberg, M. and Sjöqvist, F. (1992) Clin. Pharmacol. Ther. 52, 12–17.
[12] Gustafson, S., Proper, J.A., Bowie, W.E.J. and Sommer, S.S. (1987) Anal. Biochem. 165, 294–299.
[13] Umeno, M., McBride, O.W., Yang, C.S., Gelboin, H.V. and Gonzalez, F.J. (1988) Biochemistry 27, 9006–9013.
[14] Yue, Q.Y., Bertilsson, L., Puustinen-Dahl, M.-L., Säwe, J., Sjöqvist, F., Johansson, I. and Ingelman-Sundberg, M. (1989) Lancet ii (Oct. 7), 870.
[15] Johansson, I., Yue, Q.Y., Dahl, M.-L., Heim, M., Bertilsson, L., Meyer, U.A., Sjöqvist, F. and Ingelman-Sundberg, M. (1991) Eur. J. Clin. Pharmacol. 40, 553–556.
[16] Johansson, I., Porsmyr-Palmertz, M., Yue, Q.Y., Dahl, M.-L., Heim, M., Bertilsson, L., Sjöqvist, F. and Ingelman-Sundberg, M. (1992) J. Basic Clin. Physiol. Pharmacol. 3, 129.
[17] Kato, S., Shields, P.G., Caporaso, N.E., Hoover, R.N., Trump, B.F., Sugimura, H., Weston, A. and Harris, C.C. (1992) Cancer Res. 52, 6712–6715.

NOTE ADDED IN PROOF

Sequence analysis has now revealed that the polymorphic *Dra*I site in intron 6 is caused by a 7,766 A → T mutation (Hu, Y., unpublished results).